Catalog Extraction in SZ Cluster Surveys: a matched filter approach

Jean–Baptiste Melin[1,2] [*], James G. Bartlett[1], and Jacques Delabrouille[1]

[1]APC, 11 pl. Marcelin Berthelot, 75231 Paris Cedex 05, FRANCE
(UMR 7164 – CNRS, Université Paris 7, CEA, Observatoire de Paris)
[2]Department of Physics, University of California Davis, One Shields Avenue, Davis, CA , 95616 USA
e-mail: `jean-baptiste.melin@cea.fr, bartlett@apc.univ-paris7.fr, delabrouille@apc.univ-paris7.fr`

Abstract. We present a method based on matched multifrequency filters for extracting cluster catalogs from Sunyaev–Zel'dovich (SZ) surveys. We evaluate its performance in terms of completeness, contamination rate and photometric recovery for three representative types of SZ survey: a high resolution single frequency radio survey (AMI), a high resolution ground–based multiband survey (SPT), and the Planck all–sky survey. These surveys are not purely flux limited, and they loose completeness significantly before their point–source detection thresholds. Contamination remains relatively low at $< 5\%$ (less than 30%) for a detection threshold set at S/N=5 (S/N=3). We identify photometric recovery as an important source of catalog uncertainty: dispersion in recovered flux from multiband surveys is larger than the intrinsic scatter in the $Y - M$ relation predicted from hydrodynamical simulations, while photometry in the single frequency survey is seriously compromised by confusion with primary cosmic microwave background anisotropy. The latter effect implies that follow–up observations in other wavebands (e.g., 90 GHz, X–ray) of single frequency surveys will be required. Cluster morphology can cause a bias in the recovered $Y - M$ relation, but has little effect on the scatter; the bias would be removed during calibration of the relation. Point source confusion only slightly decreases multiband survey completeness; single frequency survey completeness could be significantly reduced by radio point source confusion, but this remains highly uncertain because we do not know the radio counts at the relevant flux levels.

Key words.

1. Introduction

Galaxy cluster catalogs play an important role in cosmology by furnishing unique information on the matter distribution and its evolution. Cluster catalogs, for example, efficiently trace large–scale features, such as the recently detected baryon oscillations (Eisenstein et al. 2005, Cole et al. 2005, Angulo et al. 2005, Huetsi 2005), and provide a sensitive gauge of structure growth back to high redshifts (Oukbir & Blanchard 1992, Rosati, Borgani & Norman 2002, Voit 2004 and references therein). This motivates a number of ambitious projects proposing to use large, deep catalogs to constrain both galaxy evolution models and the cosmological parameters, most notably the dark energy abundance and equation–of–state (Haiman, Mohr & Holder 2000, Weller & Battye 2003, Wang et al. 2004). Among the most promising are surveys based on the Sunyaev–Zel'dovich (SZ) effect (Sunyaev & Zeldovich 1970,

Sunyaev & Zeldovich 1972 and see Birkinshaw 1999, Carlstrom, Holder & Reese 2002 for reviews), because it does not suffer from surface brightness dimming and because we expect the observed SZ signal to tightly correlate to cluster mass (Bartlett 2001, Motl et al. 2005). Many authors have investigated the scientific potential of SZ surveys to constrain cosmology (e.g., Barbosa et al. 1996, Colafrancesco et al. 1997, Holder et al. 2000, Kneissl et al. 2001, Benson et al. 2002), emphasizing the advantages intrinsic to observing the SZ signal.

Cosmological studies demand statistically pure catalogs with well understood selection criteria. As just said, SZ surveys are intrinsically good in this light; however, many other factors – related, for example, to instrumental properties, observing conditions, astrophysical foregrounds and data reduction algorithms – influence the selection criteria. This has prompted some authors to begin more careful scrutiny of SZ survey selection functions in anticipation of future observations (Bartlett 2001, Schulz & White 2003, White 2003, Vale & White 2005, Melin et al. 2005, Juin et al. 2005).

Send offprint requests to: J.–B. Melin

[*] *New address:* CEA Saclay, DAPNIA/SPP, 91191 Gif-sur-Yvette, FRANCE

In Melin et al. (2005), we presented a general formalism for the SZ selection function together with some preliminary applications using a matched–filter cluster detection method. In this *paper* we give a thorough presentation of our cluster detection method and evaluate its performance in terms of catalog completeness, contamination and photometric recovery. We focus on three types of SZ survey: single frequency radio surveys like the Arcminute MicroKelvin Imager (AMI interferometer) survey[1], multi–band ground–based bolometric surveys such as the South Pole Telescope (SPT) survey[2], and the space–based Planck survey[3]. In each case, we quantify the selection function using the formalism of Melin et al. (2005).

We draw particular attention to the oft–neglected issue of photometry. Even if the SZ flux–mass relation is intrinsically tight, what matters in practice is the relation between the *observed* SZ flux and the mass. Photometric errors introduce both bias and additional scatter in the observed relation. Calibration of the $Y - M$ relation will in principal remove the bias; calibration precision, however, depends crucially on the scatter in the observed relation. Good photometry is therefore very important. As we will see, observational uncertainty dominates the predicted intrinsic scatter in this relation in all cases studied.

We proceed as follows. In section 2, we discuss cluster detection techniques and present the matched filter formalism. We describe our detection algorithm in Section 3. Using Monte Carlo simulations of the three types of survey, we discuss catalog completeness, contamination and photometry. This is done in Section 4 under the ideal situation where the filter perfectly matches the simulated clusters and in the absence of point sources. In Section 5 we examine effects caused by cluster morphology, using N–body simulations, and then the effect of point sources. We close with a final discussion and conclusions in Section 6.

2. Detecting Clusters

The detection and photometry of extended sources presents a complexity well appreciated in Astronomy. Many powerful algorithms, such as *SExtractor* (Bertin & Arnouts 1996), have been developed to extract extended sources superimposed on an unwanted background. They typically estimate the local background level and group pixels brighter than this level into individual objects. Searching for clusters at millimeter wavelengths poses a particular challenge to this approach, because the clusters are embedded in the highly variable background of the primary CMB anisotropies and Galactic emission. Realizing the importance of this issue, several authors have proposed specialized techniques for SZ cluster detection. Before detailing our own method,

we first briefly summarize some of this work in order to motivate our own approach and place it in context.

2.1. Existing Algorithms

Diego et al. (2002) developed a method designed for the Planck mission that is based on application of *SExtractor* to SZ signal maps constructed by combining different frequency channels. It makes no assumption about the frequency dependance of the different astrophysical signals, nor the cluster SZ emission profile. The method, however, requires many low–noise maps over a broad range of frequencies in order to construct the SZ map to be processed by *SExtractor*. Although they will benefit from higher resolution, planned ground–based surveys will have fewer frequencies and higher noise levels, making application of this method difficult.

In another approach, Herranz et al. (2002a, 2002b; see also López-Caniego et al. 2005 for point–source applications) developed an ingenious filter (*Scale Adaptive Filter*) that simultaneously extracts cluster size and flux. Defined as the optimal filter for a map containing a single cluster, it does not account for source blending. Cluster–cluster blending could be an important source of confusion in future ground–based experiments, with as a consequence poorly estimated source size and flux.

Hobson & McLachlan (2003) recently proposed a powerful Bayesian detection method using a Monte Carlo Markov Chain. The method simultaneously solves for the position, size, flux and morphology of clusters in a given map. Its complexity and run–time, however, rapidly increase with the number of sources.

More recently, Schäfer et al. (2004) generalized scale adaptive and matched filters to the sphere for the Planck all–sky SZ survey. Pierpaoli et al. (2004) propose a method based on wavelet filtering, studying clusters with complex shapes. Vale & White (2005) examine cluster detection using different filters (matched, wavelets, mexican hat), comparing completeness and contamination levels.

Finally, Pires et al. (2005) introduced an independent component analysis on simulated multi–band data to separate the SZ signal, followed by non–linear wavelet filtering and application of SExtractor.

Our aim is here is two–fold: to present and extensively evaluate our own SZ cluster catalog extraction method, and to use it in a comprehensive study of SZ survey selection effects. The two are in fact inseparable. First of all, selection effects are specific to a particular catalog extraction method. Secondly, we require a robust, rapid algorithm that we can run over a large number of simulated data sets in order to accurately quantify the selection effects. This important consideration conditions the kind of extraction algorithm that we can use. With this in mind, we have developed a fast catalog construction algorithm based on matched filters for both single and multiple frequency surveys. It is based on the approach first proposed by Herranz et al., but accounts for source blending.

---

[1] http://www.mrao.cam.ac.uk/telescopes/ami/

[2] http://astro.uchicago.edu/spt/
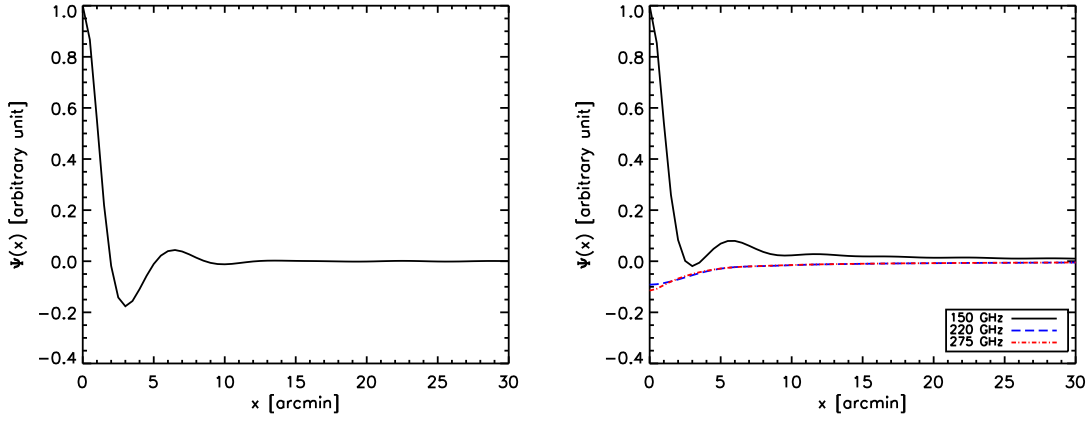
[3] http://astro.estec.esa.nl/Planck/

Fig. 1. Two examples of the matched filter for $\theta_c = 1$ arcmin. The curves give the radial profiles of the filters, which are symmetric because we have chosen a symmetric cluster template. *Left:* filter for a single frequency survey with a $\theta_{fwhm} = 1.5$ arcmin beam and 8 μK instrumental noise/beam (AMI–like, see Table 1). The undulating form of the filter maximizes the cluster signal while reducing contamination from primary CMB anisotropy. *Right:* The three components of the 3–band filter for a SPT–like experiment (Table 1). The filter is arbitrarily normalized to unity at 150 GHz. The filter uses both spatial and frequency weighting to optimally extract the cluster signal from the CMB and instrument noise. Although in this Figure the filters continue to large radii, in practice we truncate them at $10\theta_c$.

After describing the method, we apply the formalism given in Melin et al. (2005) to quantify the selection function and contamination level in up–coming SZ surveys. We take as representative survey configurations AMI, SPT and Planck, and Monte Carlo simulate the entire catalog extraction process from a large ensemble of realizations for each configuration. By comparing to the simulated input catalogs, we evaluate the extracted catalogs in terms of their completeness, contamination and photometric accuracy/precision. We will place particular emphasis on the importance of the latter, something which has received little attention in most studies of this kind.

2.2. Matched Filters

The SZ effect is caused by the hot gas ($T \sim 1 - 10$ keV) contained in galaxy clusters known as the intracluster medium (ICM); electrons in this gas up–scatter CMB photons and create a unique spectral distortion that is negative at radio wavelengths and positive in the submillimeter, with a zero–crossing near 220 GHz. The form of this distortion is universal (in the non–relativistic limit applicable to most clusters), while the amplitude is given by the Compton y parameter, an integral of the gas pressure along the line–of–sight. In a SZ survey, clusters will appear as sources extended over arcminute scales (apart from the very nearby objects, which are already known) with brightness profile

$$\Delta i_\nu(\boldsymbol{x}) = y(\boldsymbol{x})j_\nu \tag{1}$$

relative to the mean CMB brightness. Here $y(\boldsymbol{x})$ is the Compton y parameter at position $\boldsymbol{x}$ (a 2D vector on the sky) and j_ν is the SZ spectral function evaluated at the observation frequency ν.

Matched filters for SZ observations were first proposed by Haehnelt & Tegmark (1996) as a tool to estimate cluster peculiar velocities from the kinetic effect, and Herranz et al. (2002a, 2002b) later showed how to use them to detect clusters via the thermal SZ effect. They are designed to maximally enhance the signal–to–noise for a SZ cluster source by optimally (in the least square sense) filtering the data, which in our case is a sky map or set of maps at different frequencies. They do so by incorporating prior knowledge of the cluster signal, such as its spatial and spectral characteristics. The unique and universal frequency spectrum of the thermal SZ effect (in the non–relativistic regime) is hence well suited for a matched–filter approach.

Less clear is the choice of the spatial profile $T_{\theta_c}(\boldsymbol{x})$ to adopt for cluster SZ emission. One aims to choose a spatial template that represents as well as possible the average SZ emission profile. In other words, we want $T_{\theta_c}(\boldsymbol{x}) = \langle y(\boldsymbol{x})/y_o \rangle_C$, where the average is over many clusters of size θ_c. In the following, we choose to describe clusters with a projected spherical β–profile:

$$y(\boldsymbol{x}) = y_o(1 + |\boldsymbol{x}|^2/\theta_c^2)^{-(3\beta-1)/2} \tag{2}$$

with $\beta = 2/3$ (with one exception, shown for comparison in Figure 2). The spatial template is therefore described by a single parameter, the core radius θ_c; in our calculations, we truncate the profile at $10\theta_c$. This is a reasonable choice, given X–ray observations (Arnaud 2005) of the intracluster medium and the resolution of planned SZ surveys.

In reality, of course, we know neither this average profile precisely nor the dispersion of individual clusters around it beforehand. This is an important point, because our choice for the template will affect both the detection efficiency and photometric accuracy. Detection efficiency

will be reduced if the template does not well represent the average profile and, as will become clear below, the photometry will be biased. In general, *the survey selection function unavoidably suffers from uncertainty induced by unknown source astrophysics* (in addition to other sources of uncertainty).

In the following, we first study (Section 4) the ideal case where the filters perfectly match the cluster profiles, i.e., we use the β–model for both our simulations and as the detection template. In a later section (5), we examine the effects caused by non–trivial cluster morphology, as well as by point source confusion.

Consider a cluster with core radius θ_c and central y–value y_o positioned at an arbitrary point $\boldsymbol{x}_o$ on the sky. For generality, suppose that the region is covered by several maps $M_i(\boldsymbol{x})$ at N different frequencies ν_i $(i = 1, ..., N)$. We arrange the survey maps into a column vector $\boldsymbol{M}(\boldsymbol{x})$ whose i^{th} component is the the map at frequency ν_i; this vector reduces to a scalar map in the case of a single frequency survey. Our maps contain the cluster SZ signal plus noise:

$$\boldsymbol{M}(\boldsymbol{x}) = y_o \boldsymbol{j}_\nu T_{\theta_c}(\boldsymbol{x} - \boldsymbol{x}_o) + \boldsymbol{N}(\boldsymbol{x}) \tag{3}$$

where $\boldsymbol{N}$ is the noise vector (whose components are noise maps at the different observation frequencies) and $\boldsymbol{j}_\nu$ is a vector with components given by the SZ spectral function j_ν evaluated at each frequency. Noise in this context refers to both instrumental noise as well as all signals other than the cluster thermal SZ effect; it thus also comprises astrophysical foregrounds, for example, the primary CMB anisotropy, diffuse Galactic emission and extragalactic point sources.

We now build a filter $\boldsymbol{\Psi}_{\theta_c}(\boldsymbol{x})$ (in general, a column vector in frequency space) that returns an estimate, $\hat{y}_o$, of y_o when centered on the cluster:

$$\hat{y}_o = \int d^2 x \; \boldsymbol{\Psi}_{\theta_c}{}^t(\boldsymbol{x} - \boldsymbol{x}_o) \cdot \boldsymbol{M}(\boldsymbol{x}) \tag{4}$$

where superscript t indicates a transpose (with complex conjugation when necessary). This is just a linear combination of the maps, each convolved with its frequency–specific filter $(\Psi_{\theta_c})_i$. We require an unbiased estimate of the central y value, so that $\langle \hat{y}_o \rangle = y_o$, where the average here is over both total noise and cluster (of core radius θ_c) ensembles. Building the filter with the known SZ spectral form and adopted spatial template optimizes the signal–to–noise of the estimate; in other words, the filter is *matched* to the prior information. The filter is now uniquely specified by demanding a minimum variance estimate. The result expressed in Fourier space (the flat sky approximation is reasonable on cluster angular scales) is (Haehnelt & Tegmark 1996, Herranz et al. 2002a, Melin et al. 2005):

$$\boldsymbol{\Psi}_{\theta_c}(\boldsymbol{k}) = \sigma_{\theta_c}^2 \boldsymbol{P}^{-1}(\boldsymbol{k}) \cdot \boldsymbol{F}_{\theta_c}(\boldsymbol{k}) \tag{5}$$

where

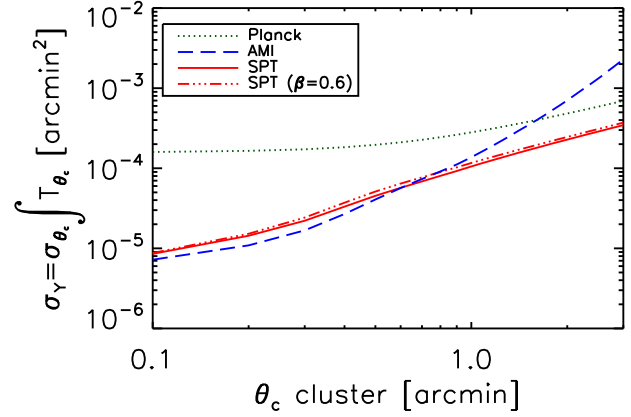$$\boldsymbol{F}_{\theta_c}(\boldsymbol{k}) \equiv \boldsymbol{j}_\nu T_{\theta_c}(\boldsymbol{k}) \tag{6}$$

Fig. 2. Filter noise expressed in terms of integrated SZ flux Y – $\sigma_Y = \sigma_{\theta_c} \int T_{\theta_c}(\boldsymbol{x}) \, d\boldsymbol{x}$ – as a function of template core radius θ_c for the three experiments listed in Table 1. A cluster with $Y = \sigma_Y$ would be detected at a signal–to–noise ratio $q = 1$. At a fixed detection threshold q (e.g., 3 or 5), the completeness of a survey rapidly increases from zero to unity in the region above its corresponding curve $q\sigma_Y(\theta_c)$ (Melin et al. 2005). All the curves adopt our fiducial value of $\beta = 2/3$, except the dashed–triple–dotted red curve, shown for comparison, which corresponds to the SPT case with $\beta = 0.6$; this curve is systematically higher by (2.5 to 13)%, depending on θ_c.

$$\sigma_{\theta_c} \equiv \left[\int d^2 k \; \boldsymbol{F}_{\theta_c}{}^t(\boldsymbol{k}) \cdot \boldsymbol{P}^{-1} \cdot \boldsymbol{F}_{\theta_c}(\boldsymbol{k}) \right]^{-1/2} \tag{7}$$

with $\boldsymbol{P}(\boldsymbol{k})$ being the noise power spectrum, a matrix in frequency space with components P_{ij} defined by $\langle N_i(\boldsymbol{k}) N_j^*(\boldsymbol{k}') \rangle_N = P_{ij}(\boldsymbol{k}) \delta(\boldsymbol{k} - \boldsymbol{k}')$. The quantity σ_{θ_c} gives the total noise variance through the filter. When we speak of the signal–to–noise of a detection, we refer to $\hat{y}_o / \sigma_{\theta_c}$.

We write the noise power spectrum as a sum $P_{ij} = P_i^{\text{noise}} \delta_{ij} + B_i(\boldsymbol{k}) B_j^*(\boldsymbol{k}) P_{ij}^{\text{sky}}$, where P_i^{noise} represents the instrumental noise power in band i, $B(\boldsymbol{k})$ the observational beam and P_{ij}^{sky} gives the foreground power (non–SZ signal) between channels i and j. As explicitly written, we assume uncorrelated instrumental noise between observation frequencies. Note that we treat the astrophysical foregrounds as isotropic, stationary random fields with zero mean. The zero mode is, in any case, removed from each of the maps, and the model certainly applies to the primary CMB anisotropy. It should also be a reasonable model for fluctuations of other foregrounds about their mean, at least over cluster scales[4].

Two examples of the matched filter for $\theta_c = 1$ arcmin are shown in Fig. 1, one for an AMI–like single frequency survey with a 1.5 arcmin beam (left–hand panel) and the other for a SPT–like 3–band filter (right–hand panel); see

---

[4] We make no assumption about the Gaussianity of the fields; the estimator remains unbiased even if they are not Gaussian, although optimality must be redefined in this case.

Table 1 for the experimental characteristics. The filters are circularly symmetric, with the figures giving their radial profiles, because we have chosen a spherical cluster model. We clearly see the spatial weighting used by the single frequency filter to optimally extract the cluster from the noise and CMB backgrounds. The multiple frequency filter Ψ_{θ_c} is a 3–element column vector containing filters for each individual frequency. In this case, the filter employs both spectral and spatial weighting to optimally extract the cluster signal.

Figure 2 shows the filter noise as a function of template core radius θ_c. We plot the filter noise expressed in terms of an equivalent noise $\sigma_Y \equiv \sigma_{\theta_c} \int T_{\theta_c}(\boldsymbol{x}) \, d\boldsymbol{x}$ on the integrated SZ flux Y. The dashed–triple–dotted red curve with $\beta = 0.6$ is shown for comparison to gauge the impact of changing this parameter, otherwise fixed at $\beta = 2/3$ throughout this work. Melin et al. (2005) use the information in this figure to construct survey completeness functions. At fixed signal–to–noise q, the completeness of a survey rapidly increases to unity in the region above the curve $q\sigma_Y$. The Figure shows that high angular resolution ground–based surveys (e.g., AMI, SPT) are not purely flux limited, because their noise level rises significantly with core radius. The lower resolution of the Planck survey, on the other hand, results in more nearly flux limited sample.

3. Catalog Extraction

Catalog construction proceeds in three steps, the last two of which are repeated[5]:

1. Convolution of the frequency map(s) with matched filters corresponding to different cluster sizes;
2. Identification of candidate clusters as objects with signal–to–noise $\hat{y}_o/\sigma_{\theta_c} > q$, where q is our fixed detection threshold, followed by photometry of the brightest remaining cluster candidate, which is then added to the final cluster catalog;
3. Removal of this object from the set of filtered maps using the photometric parameters (e.g., y_o and θ_c) from the previous step.

We loop over the last two steps until there are no remaining candidates above the detection threshold. The following sections detail each step.

3.1. Map filtering

In the first step, we convolve the observed map(s) with matched filters covering the expected range of

---

[5] Note that we have made some changes in the two last steps compared to the description given in Melin et al. (2005). We no longer sort candidates in a tree structure for de-blending; instead, we identify and then remove candidates one by one from the filtered maps. This has only a small impact on the completeness of the detection algorithm, leaving the conclusions of our previous paper intact. The changes, however, greatly improve photometry and lower contamination.

core radii. For AMI and SPT, for example, we vary θ_c from 0.1 to 3 arcmins in 0.1 steps (i.e., $\theta_c = 0.1, 0.2, ..., 2.9, 3$ arcmins) and add three values for the largest clusters: $4, 5, 6$ arcmins. We thus filter the map(s) n_{θ_c} times ($n_{\theta_c} = 33$ for AMI and SPT) to obtain $2\, n_{\theta_c}$ filtered maps, J_{θ_c} et L_{θ_c}. The n_{θ_c} maps J_{θ_c} give the SZ amplitude (obtained using Ψ_{θ_c}), while the n_{θ_c} maps L_{θ_c} give the signal–to–noise ratio: $L_{\theta_c} = J_{\theta_c}/\sigma_{\theta_c}$). We set a detection threshold at fixed signal–to–noise q and identify candidates at each filter scale θ_c as pixels with $L_{\theta_c} > q$. Common values for the threshold are $q = 3$ and $q = 5$; the choice is a tradeoff between detection and contamination rates (see below).

3.2. Cluster parameter estimation: Photometry

We begin the second step by looking for the brightest candidate pixel in the set of maps L_{θ_c}. The candidate cluster is assigned the spatial coordinates (x, y) of this pixel, and its core radius is defined as the filter scale of the map containing the pixel: $\theta_c = \theta_f$. We then calculate the total integrated flux using $Y = \hat{y}_o \int T_{\theta_c}(\boldsymbol{x}) \, d\boldsymbol{x}$, where $\hat{y}_o$ is taken from the map J_{θ_c} at the same filter scale. We refer to this step as the photometric step, and the parameters $\hat{y}_o$, θ_c and Y as photometric parameters. Note that measurement error on Y comes from errors on both $\hat{y}_o$ and θ_c (We return to this in greater detail in Section 4.4).

3.3. Catalog construction

The candidate cluster is now added to the final cluster catalog, and we proceed by removing it from the set of filtered maps J_{θ_c} and L_{θ_c} before returning to step 2. To this end, we construct beforehand a 2D array (library) of un–normalized, filtered cluster templates (postage–stamp maps)

$$\mathcal{T}_{\theta_c, \theta_f}(\boldsymbol{x}) = \int d^2x' \ \Psi_{\theta_f}(\boldsymbol{x}' - \boldsymbol{x}) T_{\theta_c}(\boldsymbol{x}') \tag{8}$$

with the cluster centered in the map. Note that θ_c runs over core radius and θ_f over filter scale. At each filter scale θ_f, we place the normalized template $\hat{y}_o \mathcal{T}_{\theta_c, \theta_f}$ on the cluster position (x, y) and subtract it from the map. The library of filtered templates allows us to perform this step rapidly.

We then return to step 2 and repeat the process until there are no remaining candidate pixels. Thus, clusters are added to the catalog while being subtracted from the maps one at a time, thereby de-blending the sources. By pulling off the brightest clusters first, we aim to minimize uncertainty in the catalog photometric parameters. Nevertheless, it must be emphasized that the entire procedure relies heavily on the use of templates and that real clusters need not conform to the chosen profiles. We return to the effects of cluster morphology below.

In the end, we have a cluster catalog with positions (x, y), central Compton y parameters, sizes θ_c and fluxes Y.

Type	Frequencies [GHz]	Res. *fwhm* [arcmin]	Inst. noise [μK/beam]	Area [deg^2]
AMI	15	1.5	8	10
SPT	150	1	10	
	220	0.7	60	4000
	275	0.6	100	
Planck	143	7.1	6	
	217	5	13	41253
	353	5	40	

Table 1. Characteristics of the three types of experiments considered. We run our extraction method on 100 sky patches of 3×3 square degrees (for AMI and SPT) and 12×12 square degrees (for Planck).

4. Cluster recovery

We tested our catalog construction method on simulated observations of the three representative types of SZ survey specified in Table 1. The simulations include SZ emission, primary CMB anisotropy and instrumental noise and beam smearing. We do not include diffuse Galactic foregrounds in this study. We begin in this section with the ideal case where the filter perfectly matches the simulated clusters (spherical β–model profiles) and in the absence of extragalactic point sources. We return to the additional effects of cluster morphology and point source confusion in Section 5.

The simulated maps are generated by Monte Carlo. We first create a realization of the linear matter distribution in a large box using the matter power spectrum. Clusters are then distributed according to their expected number density, given by the mass function, and bias as a function of mass and redshift. We also give each cluster a peculiar velocity consistent with the matter distribution according to linear theory. The simulations thus featuring cluster spatial and velocity correlations accurate first order, which is a reasonable approximation on cluster scales. In this paper, we use these simulations but we do not study the impact of the correlations on the detection method, leaving this issue to forthcoming work.

The cluster gas is modeled by a spherical isothermal β–profile with $\beta = 2/3$ and $\theta_c/\theta_v = 0.1$, where θ_v is the angular projection of the virial radius and which varies with cluster mass and redshift following a self-similar relationship. We choose an $M - T$ relation consistent with the local abundance of X–ray clusters and our value of σ_8, given below (Pierpaoli et al. 2004). Finally, we fix the gas mass fraction at $f_{gas} = 0.12$ (e.g., Mohr et al. 1999). The input catalog consists of clusters with total mass $M > 10^{14} M_\odot$, which is sufficient given the experimental characteristics listed in Table 1. Delabrouille et al. (2002) describe the simulation method in more detail.

We generate primary CMB anisotropies using the power spectrum calculated by CMBFAST[6] (Seljak & Zaldarriaga 1996) for a flat concordance model

[6] http://cmbfast.org/

with $\Omega_M = 0.3 = 1 - \Omega_\Lambda$ (Spergel et al. 2003), Hubble constant $H_o = 70$ km/s/Mpc (Freedman et al. 2001) and a power spectrum normalization $\sigma_8 = 0.98$. As a last step we smooth the map with a Gaussian beam and add Gaussian white noise to model instrumental effects[7].

We simulate maps that would be obtained from the proposed surveys listed in Table 1. The first is an AMI[8]–like experiment (Jones et al. 2002), a single frequency, high resolution interferometer; the sensitivity corresponds to a one–month integration time per 0.1 square degree (Kneissl et al. 2001). The SPT[9]–like experiment (Ruhl et al. 2004) is a high resolution, multi–band bolometer array. We calculate the noise levels assuming an integration time of 1 hour per square degree, and a split of 2/3, 1/6, 1/6 of the 150, 220, 275 GHz channels for the 1000 detectors in the focal plane array (Ruhl et al. 2004). Finally, we consider the space–based Planck[10]–like experiment, with a nominal sensitivity for a 14 month mission. For the AMI and SPT maps we use pixels[11] of 30 arcsec, while for Planck the pixels are 2.5 arcmin.

We simulate 100 sky patches of 3×3 square degrees for both AMI and SPT, and of 12×12 square degrees for Planck. This is appropriate given the masses of detected clusters in each experiment. In practice, AMI will cover a few square degrees, similar to the simulated patch, while SPT will cover 4000 square degrees and Planck will observe the entire sky. Thus, the surveys decrease in sensitivity while increasing sky coverage from top to bottom in Table 2 (see also Table 1).

deg^{-2}	$S/N > 3$	$S/N > 5$
AMI	44	20
	(38)	(16)
SPT	35	12
	(27)	(11)
Planck	1.00	0.38
	(0.84)	(0.35)

Table 2. Extracted counts/sq. deg. from simulations of the three types of survey. The numbers in parenthesis give the counts predicted by our analytic cluster model; the difference is due to cluster overlap confusion (see text).

[7] The 3–year WMAP results, published after the work presented here was finished, favor a significantly lower value of σ_8 (Spergel et al. 2006). This could lower the total number of clusters in our simulations by up to a factor of ~ 2. As we are interested here in catalog recovery, where we compare output to input catalogs, this change should only cause relatively minor changes to our final results.

[8] http://www.mrao.cam.ac.uk/telescopes/ami/index.html

[9] http://astro.uchicago.edu/spt/

[10] http://www.rssd.esa.int/index.php?project=PLANCK

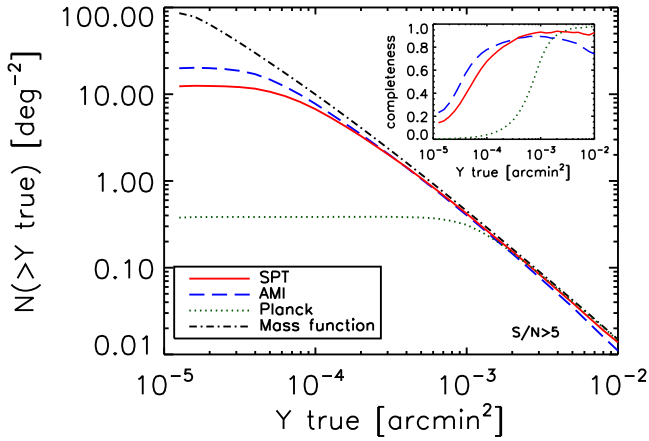[11] Pixel sizes are at least 2 times smaller than the best channel of each experiment.

Fig. 3. Cluster counts $N(> Y)$ per square degree as a function of true SZ flux Y for a threshold of $S/N > 5$. The dash–dotted black line gives the cluster counts from the mass function (Jenkins et al. 2001). The dashed blue line gives the recovered cluster counts for AMI, the red solid line for SPT and the dotted green line for Planck. The inset shows the completeness ratio (relative to the mass function prediction) for each survey. All the surveys are significantly incomplete at their point–source sensitivities (5 times the y–intercept in Figure 2).

4.1. Association criteria

An important issue for catalog evaluation is the association between a detected object (candidate cluster) with a cluster from the simulation input catalog (real cluster); in other words, a candidate corresponds to which, if any, real cluster. Any association method will be imprecise, and estimates of catalog completeness, contamination and photometric accuracy will unavoidably depend on the choice of association criteria.

We proceed as follows: for each detection, we look at all input clusters with centers positioned within a distance $r = \sqrt{8} \times d$, where d is the pixel size ($d = 30$ arcsec for AMI and SPT, $d = 2.5$ arcmin for Planck); this covers the neighboring 24 pixels. If there is no input cluster, then we have a false detection; otherwise, we identify the candidate with the cluster whose flux is closest to that of the detection. After running through all the candidates in this fashion, we may find that different candidates are associated with the same input cluster. In this case, we only keep the candidate whose flux is closest to the common input cluster, and we flag the other candidates as false detections (multiple detections).

At this stage, some associations may nevertheless be chance alignments. We therefore employ a second parameter, Y_{cut}: a candidate associated with a real cluster of flux $Y < Y_{cut}$ is flagged as a false detection. We indicate these false detections as diamonds in Figures 7, 8, 9 and 11. The idea is that such clusters are too faint to have been detected and the association is therefore by chance. In the following, we take $Y_{cut} = 1.5 \times 10^{-5}$ arcmin^2 for

AMI and SPT, respectively, and $Y_{cut} = 3 \times 10^{-4}$ arcmin^2 for Planck. Note that these numbers are well below the point–source sensitivity (at S/N=5) in each case (see below and Figure 2).

4.2. Completeness

Figure 3 shows completeness for the three experiments in terms of *true* integrated Y, while Table 2 summarizes the counts. In Figure 4 we give the corresponding limiting mass as a function of redshift. Given our cluster model, AMI, SPT and Planck should find, respectively, about 16, 11 and 0.35 clusters/deg.^{2} at a $S/N > 5$; these are the numbers given in parentheses in Table 2. Cluster overlap confusion accounts for the fact that the actual counts extracted from the simulated surveys are higher: some clusters that would not otherwise pass the detection cut enter the catalog because the filter adds in flux from close neighbors.

A detection threshold of $S/N = 5$ corresponds to a point–source sensitivity of just below $Y = 5 \times 10^{-5}$ arcmin^2 for both AMI and SPT, as can be read off the left–hand–side of Figure 2. The surveys approach a high level of completeness only at $Y > 10^{-4}$ arcmin^2, however, due to the rise of the selection cut with core radius seen in Figure 2. For these high resolution surveys, point–source sensitivity gives a false idea of the survey completeness flux limit.

At the same signal–to–noise threshold, Planck is essentially complete above $Y \sim 10^{-3}$ arcmin^2 and should detect about 0.4 clusters per square degree. Since most clusters are unresolved by Planck, the survey reaches a high completeness level near the point–source sensitivity. We also see this from the small slope of the Planck selection cut in Figure 2.

We emphasize that the surveys (in particular, the high resolution surveys) are not flux limited for any value of q, because increasing q simply translates the curve in Figure 2 along the y axis. However, one can approach a flux–limited catalog by selecting clusters at $S/N > q$ and then cutting the resulting catalog at $Y_o > Y_{limit} \equiv Q\sigma_Y(\theta_c = 0.1 \text{ arcmin})$, where the constant $Q > q$. As Q increases we tend towards a catalog for which $Y \sim Y_o > Y_{limit}$. In the case of SPT with $q = 3$, for example, we find that large values of Q (> 10) are required to approach a reasonable flux–limited catalog; this construction, however, throws away a very large number of detected clusters.

Although the AMI (single frequency) and SPT (multiband) survey maps have comparable depth, SPT will cover ~ 4000 sq. degrees, compared to AMI's ~ 10 sq. degrees. Planck will only find the brightest clusters, but with full sky coverage. Predictions for the counts suffer from cluster modeling uncertainties, but the comparison between experiments is robust and of primary interest here.
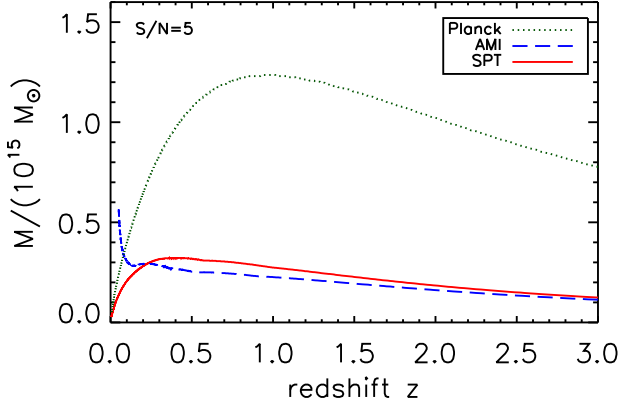
Fig. 4. Mininum detectable cluster mass as a function of redshift, $M(z)$, corresponding to $S/N = 5$ for the three experiments discussed in the text. The rise at low redshift for the single–frequency (AMI) curve is caused by confusion with primary CMB anisotropy.
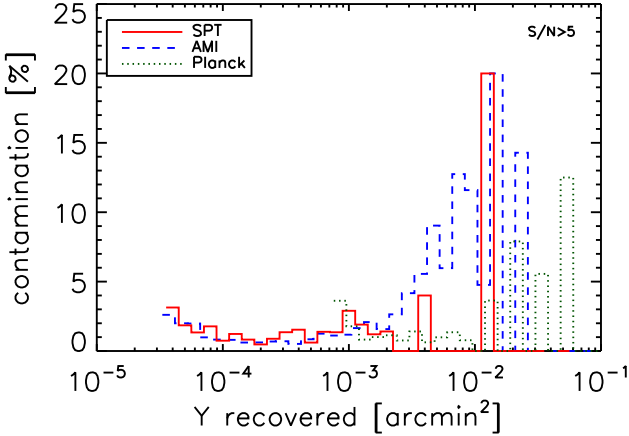

Fig. 5. Contamination as a function of the core radius θ_c for the three experiments and for $S/N > 5$.

4.3. Contamination

Figure 5 shows the contamination level at $S/N > 5$ for each survey type as a function of *recovered* flux Y_o. The multiband experiments (SPT and Planck) benefit from low contamination at all fluxes. Single frequency surveys (e.g., AMI), on the other hand, experience a slightly higher contamination level at large flux due to confusion from primary CMB anisotropy. This confusion also degrades the photometry, as we discuss below.

At $S/N > 5$, the AMI, SPT and Planck catalogs have less than 2% total contamination rate. These numbers increase to $\sim$ 23, 20 and 27 percent, respectively, for AMI, SPT and Planck at a detection threshold of $S/N > 3$. Note that the total contamination rate is an average over the histogram of Figure 5 weighted by the number of ob-
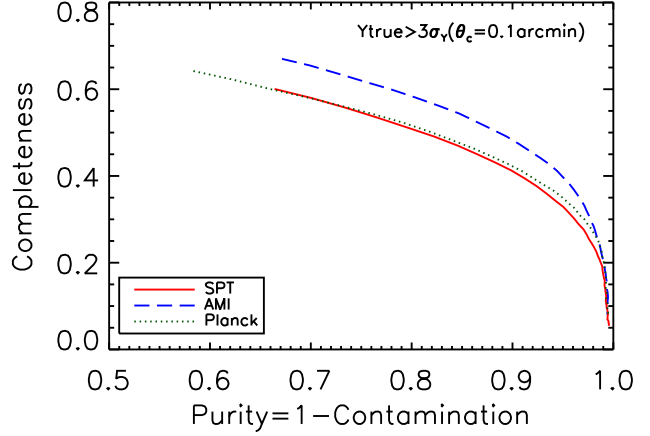

Fig. 6. Completeness-Purity plot. For each curve, q varies from 3 (top-left) to 10 (bottom-right). For each experiment, the input catalog contains clusters with true flux greater than three times the point source sensitivity (Ytrue $> 2.2 \times 10^{-5}$ arcmin^2 for AMI, Ytrue $> 2.6 \times 10^{-5}$ arcmin^2 for SPT and Ytrue $> 4.8 \times 10^{-4}$ arcmin^2 for Planck). See text for details.
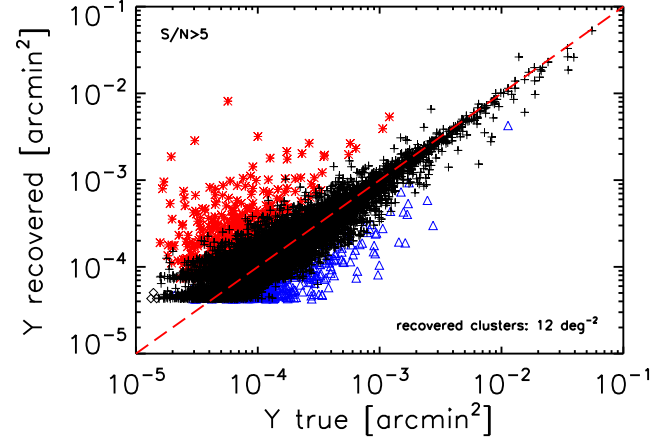

Fig. 7. Recovered vs. true flux for SPT clusters extracted at $S/N > 5$ from 100 survey simulations. The diamonds indicate cluster detections with $Y < Y_{cut}$, which we take as false detections. The mean trend $Y_o(Y)$ has a slight bias (see text) and a roughly constant scatter of $\sigma_{logY_o} = 0.17$ over the interval in true Y from 10^{-4} arcmin^2 to 4×10^{-3} arcmin^2. The clusters which have their core radii overestimated by a factor of 2 are plotted as red crosses and the clusters which have their core radii underestimated by a factor of 2 are plotted as blue triangles.

jects in each bin; thus, the higher contamination at large flux is down–weighted in the total rate.

In all cases, the contamination rate is higher than expected from pure Gaussian noise fluctuations; there is an important contribution from cluster–cluster confusion (residuals from cluster subtraction and overlaps). We expect even higher contamination rates in practice, because

of variations in cluster morphology around the filter templates. We quantify this latter effect below.

A useful summary of these results is a completeness–purity plot, as shown in Figure 6. Proper comparison of the different experiments requires an appropriate choice of input catalog used to define the completeness in this plot. Here, we take the input catalog as all clusters with (true) flux geater than three times the point source sensitivity for each experiment. If the clusters were point sources and the detection method perfect (i.e. not affected by confusion), the completeness would be 1 for $q = 3$ in the top-left corner. These curves summarize the efficiency of our cluster detection method; however, they give no information on the photometric capabilities of the experiments.

4.4. Photometry

We now turn to the important, but often neglected issue of cluster SZ photometry. The ability of a SZ survey to constrain cosmology relies on application of the $Y - M$ relation. As mentioned, we expect the *intrinsic* (or *true*) flux to tightly correlate with cluster mass (Bartlett 2001), as indeed borne out by numerical simulations (da Silva et al. 2004, Motl et al. 2005, Nagai 2005). Nevertheless, unknown cluster physics could affect the exact form and normalization of the relation, pointing up the necessity of an empirical calibration (referred to as survey *calibration*), either with the survey data itself (self–calibration, Hu 2003, Majumdar & Mohr 2003, Lima & Hu 2004, Lima & Hu 2005) or using external data, such as lensing mass estimates (Bartelmann 2001) (although the latter will be limited to relatively low redshifts).

Photometric measurement accuracy and precision is as important as cluster physics in this context: what matters in practice is the relation between *recovered* SZ flux Y_o and cluster mass M. Biased SZ photometry (bias in the $Y - Y_o$) relation will change the form and normalization of the $Y_o - M$ relation and noise will increase the scatter. One potentially important source of photometric error for the matched filter comes from cluster morphology, i.e., the fact that cluster profiles do not exactly follow the filter shape (see Section 5).

Survey calibration will help remove the bias, but with an ease that depends on the photometric scatter: large scatter will increase calibration uncertainty and/or necessitate a larger amount of external data. In addition, scatter will degrade the final cosmological constraints (e.g., Lima & Hu 2005). Photometry should therefore be considered an important evaluation criteria for cluster catalog extraction methods.

Consider, first, SPT photometry. Figure 7 shows the relation between observed (or recovered) flux Y_o and true flux Y for a detection threshold of $S/N > 5$. Fitting for the average trend of Y_o as a function of Y, we obtain

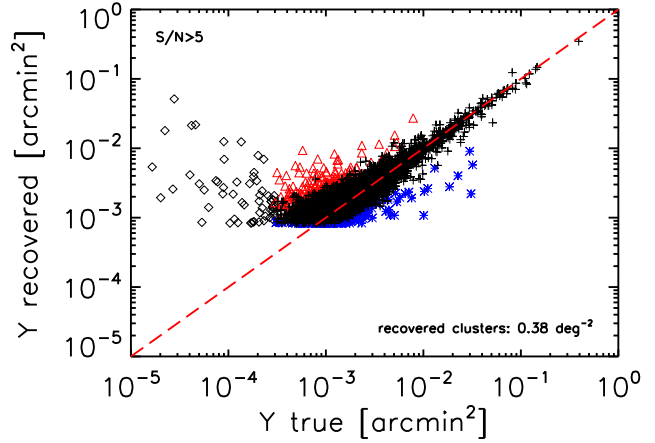$$\log Y_o = 0.96 \log Y - 0.15$$

Fig. 8. Recovered vs. true flux for Planck clusters extracted at $S/N > 5$ from 100 survey simulations. The diamonds indicate cluster detections with $Y < Y_{cut}$, which we take as false detections. The mean trend $Y_o(Y)$ has a slight bias (see text) and a roughly constant scatter of $\sigma_{\log Y_o} = 0.13$ over the interval in true Y from 2×10^{-3} arcmin^2 to 2×10^{-2} arcmin^2. The clusters which have their core radii overestimated by a factor of 2 are plotted as red crosses and the clusters which have their core radii underestimated by a factor of 2 are plotted as blue triangles.

over the interval 10^{-4} arcmin$^2 < Y < 4 \times 10^{-3}$ arcmin^2, with Y_o and Y measured in arcmin^2. There is a slight bias in that the fit deviates somewhat from the equality line, but the effect is minor. Below this flux interval, the fit curls upward in a form of Malmquist bias caused by the S/N cut (seen as the sharp lower edge on Y_o). The lack of any significant bias is understandable in this ideal case where the filter perfectly matches the cluster SZ profile. Cluster morphology, by which we mean a mismatch between the cluster SZ profile and the matched filter template), can induce bias; we return to this issue in Section 5.

The scatter about the fit is consistent with a Gaussian distribution with a roughly constant standard deviation of $\sigma_{\log Y_o} = 0.17$ over the entire interval.

The scatter is a factor of 10 larger than expected from instrumental noise alone, which is given by the selection curve in Figure 2. Uncertainty in the recovered cluster position, core radius and effects from cluster–cluster confusion all strongly influence the scatter. Photometry precision, therefore, cannot be predicted from instrumental noise properties alone, but only with simulations accounting for these other, more important effects.

Figure 8 shows the photometry for the Planck survey. Apart from some catastrophic cases (the diamonds), the photometry is good and fit by

$$\log Y_o = 0.98 \log Y - 0.07$$

over the interval 2×10^{-3} arcmin$^2 < Y < 2 \times 10^{-2}$ arcmin^2 (Y_o, Y measured in arcmin^2). The dispersion is $\sigma_{\log Y_o} = $

0.13, roughly constant over the same interval. For unre-
solved clusters, this scatter is ~ 5 times larger than the
expected instrumental–induced scatter. The brightest di-
amonds in the Figure correspond to real clusters with po-
sitional error larger than the association criteria r. As a
consequence, the candidates are falsely associated with a
small, nearby cluster, unrelated to the actual detected ob-
ject.

We emphasize that the observational scatter in the
$Y_o - Y$ relation for both SPT and Planck dominates the
intrinsic scatter of less than 5% seen in the $Y - M$ re-
lation from numerical simulations (da Silva et al. 2004,
Motl et al. 2005).

We now turn to single frequency surveys, which
Figure 9 shows to have seriously compromised photom-
etry. The distribution at a given true flux Y is in fact
bimodal, as illustrated by the solid blue histogram in
Figure 10 that gives the distribution of the recovered flux
Y_o for clusters with true flux and core radius in a bin cen-
tered on $Y = 1.5 \times 10^{-4}$ arcmin^2 and $\theta_c = 0.3$ arcmin.
We have traced this effect to an inability to accurately
determine the core radius of the candidate clusters. We
demonstrate this in Figure 11 by artificially setting the
candidate core radius to its true value taken from the asso-
ciated input cluster; the photometry now cleanly scatters
about the mean trend.

This inability to determine the core radius mainly
arises from confusion with primary CMB anisotropy, as
we now show using Figure 10. We performed 1000 simu-
lations of a single cluster ($Y = 1.5 \times 10^{-4}$ arcmin^2, $\theta_c =$
0.3 arcmin) placed at the middle of a beam–convolved
map containing background SZ clusters (from our general
simulations), primary CMB anisotropy and instrumental
noise. We then estimate its core radius and flux with our
matched filters centered on the known position (to avoid
any positional uncertainty) and trace the histogram of
resulting measured flux. This is the red dot–dashed his-
togram in the figure, which displays a bi–modality similar
to that of the blue solid histogram. We then follow the
same procedure after first removing the primary CMB
anisotropy from the simulated map. The resulting his-
togram of recovered flux is given by the black dot–dashed
line with much less pronounced bimodality. The remaining
tail reaching towards high flux is caused by cluster–cluster
confusion.

With their additional spectral information, multiband
surveys remove the primary CMB signal, thereby avoiding
this source of confusion. *The result suggests that follow–up
observations of detected clusters at a second frequency will
be required for proper photometry*; without such follow–up,
the scientific power of a single frequency survey may be
seriously compromised, as can be appreciated from inspec-
tion of Figure 9.

5. Additional Effects

As emphasized, our previous results follow for a filter that
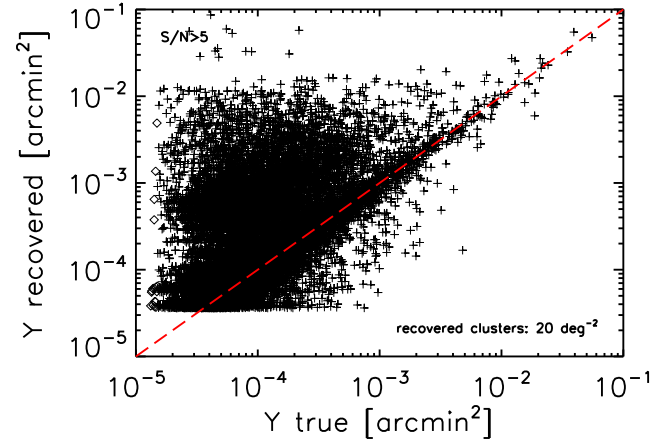perfectly matches the (spherical) clusters in our simula-

Fig. 9. Recovered vs. true flux for AMI clusters extracted
at S/N > 5 from 100 survey simulations. The diamonds
indicate cluster detections with $Y < Y_{cut}$, which we take
as false detections. The extremely large dispersion in re-
covered flux results from a bimodal distribution caused by
an inability to determine the core radius of detected clus-
ters. This inability is due to confusion from primary CMB
anisotropy, as demonstrated in Figure 10. Figure 11 shows
that reasonable photometry is possible if the core radius
can be accurately determined. This problem is specific to
single–frequency surveys that are unable to spectrally re-
move primary CMB anisotropy.

tions and in the absence of any point sources. In this sec-
tion we examine the effects of both cluster morphology
and point sources.

We find that cluster morphology has little effect on
catalog completeness, but that it does increase contami-
nation. More importantly, it can bias photometric recov-
ery, although it does not significantly increase the scatter.
This bias changes the observed $Y - M$ relation from its in-
trinsic form, adding to the modeling uncertainty already
caused by cluster gas physics. For this reason, the rela-
tion must be calibrated in order to use the SZ catalog for
any cosmological study. The observational bias would be
removed during this calibration step.

Completeness is the most affected by point source con-
fusion, decreasing somewhat for the multi–band surveys in
the presence of IR point sources. The level of confusion for
the single frequency survey remains highly uncertain due
to the unknown point source counts at low flux densities.
Contamination and photometry are essentially unaffected.

5.1. Cluster Morphology

To assess the influence of cluster morphology, we ran our
catalog extraction algorithm on maps constructed from
numerical simulations. We use the simulations presented
by Schulz & White (Schulz & White 2003) and kindly pro-
vided to us by M. White. Their simulations follow dark
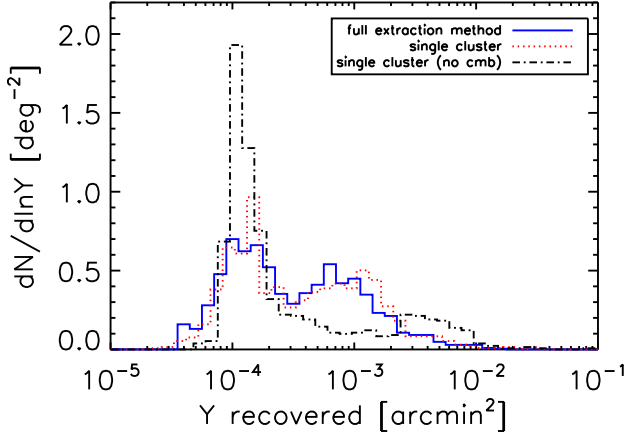matter clustering with a N–body code in a flat concor-

Fig. 10. The full blue histogram gives the cluster counts from figure 9 in the bin ($10^{-4} < Y < 2.10^{-4}$, $0.25 < \theta_c < 0.35$). We have added the cluster counts obtained from the size and flux estimation of a single cluster ($Y = 1.5 \times 10^{-4}$, $\theta_c = 0.3$) at a known position through 1000 simulations. SZ cluster background maps and the instrumental beam and noise are included. Two cases are considered : with primary CMB (dotted red histogram) and without primary CMB (dash–dotted black line). The double bump in Y recovery is visible when the primary CMB is present and disappears when it's removed showing that the primary CMB power spectrum is the cause of the double bump.
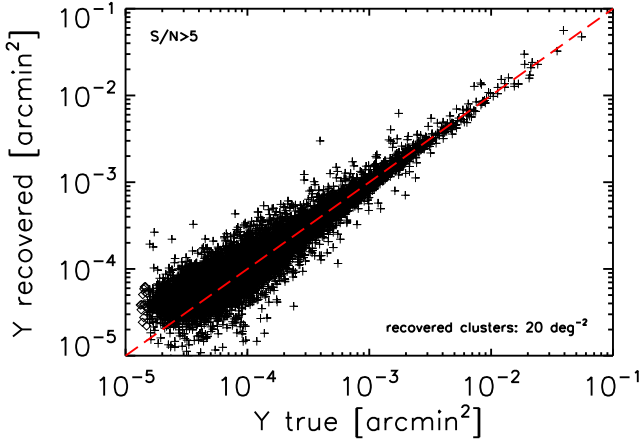

Fig. 11. Single–frequency photometry when we artificially set the core radii of detected clusters to their true values from the input catalog. The dispersion decreases dramatically, demonstrating that the inability to recover the core radius is the origin of the bad photometry seen in Figure 9.

dance cosmology, and model cluster gas physics with semi–analytical techniques by distributing an isothermal gas of mass fraction Ω_B/Ω_M according to the halo dark matter distribution. For details, see Schulz & White. In the following, we refer to these simulations as the "N–body" simulations.
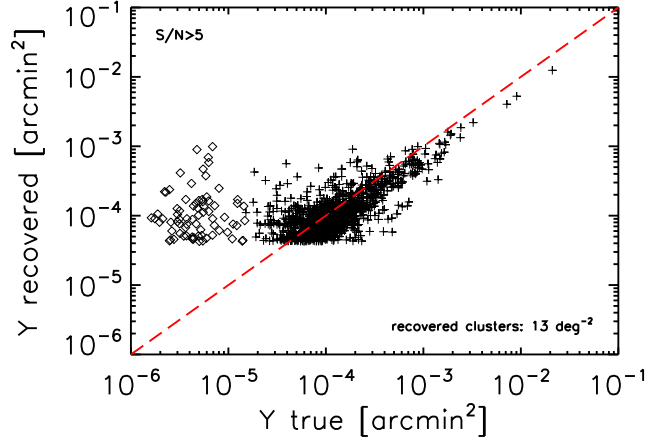

Fig. 12. Photometry for the SPT catalog from the N–body simulations. Cluster morphology (mismatch between the filter profile and the actual cluster SZ profile) clearly induces a bias between the recovered and true SZ flux. The scatter, on the other hand, is not very affected, as can be seen in comparing with Figure 7.

We proceed by comparing catalogs extracted from the N–body map to those from a corresponding simulation made with spherical clusters. The latter is constructed by applying our spherical β–model gas distribution to the cluster halos taken from the N–body simulation and using them as input to our Monte Carlo sky maps. In the process, we renormalize our $Y - M$ relation to the one used in the N–body SZ maps. We thus obtain two SZ maps containing the same cluster halos, one with spherical clusters (referred to hereafter as the "β–model" maps) and the other with more complex cluster morphology (the N-body maps). Comparison of the catalogs extracted from the two different types of simulated map gives us an indication of the sensitivity of our method to cluster morphology. We make this comparative study only for the SPT and Planck like surveys.

Catalog completeness is essentially unaffected by cluster morphology; the integrated counts, for example, follow the same curves shown in Figure 3 with very little deviation, the only difference being a very small decrease in the Planck counts at the lowest fluxes. The effect, for example, is smaller than that displayed in Figure 13 due to point source confusion (and discussed below).

Non–trivial cluster morphology, however, does significantly increase the catalog contamination rate; for example, in the SPT survey the global contamination rises from less than 2% to 13% at $S/N = 5$ for the N–body simulations. We trace this to residual flux left in the maps after cluster extraction: cluster SZ signal that deviates from the assumed spherical β–model filter profile remains in the map and is picked up later as new cluster candidates. Masking those regions where a cluster has been previously extracted (i.e., forbidding any cluster detection) drops the contamination to 4% (SPT case), but causes a decrease

of 2.8 clusters per square degree in the recovered counts; this technique would also have important consequences for clustering studies.

From Figure 12, we clearly see that cluster morphology induces a bias in the photometry. This arises from the fact that the actual cluster SZ profiles differ from the template adopted for the filter. The differences are of two types: an overall difference in the form of radial profile and local deviations about the average radial profile due to cluster substructure. It is the former that is primarily responsible for the bias. In our case, the N–body simulations have much more centrally peaked SZ emission than the filter templates, which causes the filter to systematically underestimate the total SZ flux. Cluster substructure will increase the scatter about the mean $Y_o - Y$ relation. This latter effect is not large, at least for the N–body simulations used here, as can be seen by comparing the scatter in Figures 12 and 7.

We emphasize, however, that the quantitative effects on photometry depend on the intrinsic cluster profile, and hence are subject to modeling uncertainty. The simulations used here do not include gas physics and simply assume that the gas follows the dark matter. The real bias will depend on unknown cluster physics, thus adding to the modeling uncertainty in the $Y - M$ relation. This uncertainty, due to both cluster physics and the photometric uncertainty discussed here, must be dealt with by empirically calibrating the relation, either with external data (lensing) and/or internally (self–calibration).
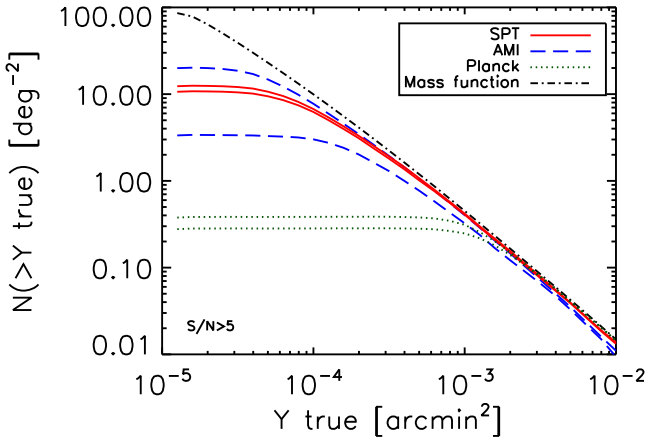

Fig. 13. Integrated cluster counts for the three types of survey. The upper curve in each pair reproduces the results of Figure 3, while the lower curve shows the effect of point source confusion. Despite the large IR point source population, multiband surveys efficiently eliminate confusion. The AMI–like survey is, on the other hand, strongly affected. This latter effect remains uncertain due to a lack of information on the faint end of the radio point source counts (see text).

5.2. Point Sources

We next examine the effect of point sources. In a previous paper (Bartlett & Melin 2005, hereafter BM) we studied their influence on survey detection sensitivity. We extend this work to our present study in this section.

Low frequency surveys, such as our AMI example, contend with an important radio source population, while higher frequency bolometer surveys face a large population of IR sources. Radio source counts down to the sub–mJy flux levels relevant for SZ surveys are unfortunately poorly known. The IR counts are somewhat better constrained at fluxes dominating the fluctuations in the IR background, although at higher frequencies (850 microns) than those used in SZ surveys; an uncertain extrapolation in frequency is thus necessary.

For the present study, we use the radio counts fit by Knox et al. 2004 to a combination of data from CBI, DASI, VSA and WMAP (see also Eq. 6 in BM), and IR counts fit to blank–field SCUBA observations at 850 microns by Borys et al. 2003 (and given by Eq. 8 in BM). We further assume that all radio sources brighter than $100\,\mu$Jy have been subtracted from our maps at 15 GHz (AMI case); this is the target sensitivity of the long baseline Ryle Telescope observations that will perform the source subtraction for AMI. No such explicit point source subtraction is readily available for the higher frequency bolometer surveys; they must rely solely on their frequency coverage to reduce point source confusion. We therefore include all IR sources in our simulations, and fix their effective spectral index $\alpha = 3$ with no dispersion[12]. We refer the reader to BM for details of our point source model. Note that for this study we use the spherical cluster model for direct comparison to our fiducial results.

Figure 13 compares the integrated counts from Figure 3 (upper curve in each case) to those extracted from the simulations including point sources (lower curves). We see that point source confusion only slightly decreases the completeness of the multiband surveys, but greatly affects the single frequency survey.

In the case of SPT, this is because point source confusion remains modest compared to the noise: the two are comparable at 150 GHz, but the noise power rises more quickly with frequency than the confusion power (see BM for details) – in other words, the noise is bluer than the confusion. This is an important consideration when looking for the optimal allocation of detectors to the observation bands.

For Planck, confusion power dominates at all frequencies, but the spectral coverage provides sufficient leverage to control it. In this light, it must be emphasized that we only include three astrophysical signals (SZ, CMB & point sources) in these simulations, so that three observation bands are sufficient. In reality, one will have to deal with other foregrounds, e.g., diffuse Galactic emission, which will require the use of additional observation bands.

---

[12] As discussed in BM, any dispersion has only a small effect on survey sensitivity

The single frequency observations, on the other hand, are strongly affected. This is consistent with the estimate in BM (Eq. 15) placing confusion noise well above instrumental noise for the chosen point source model and source subtraction threshold. We emphasize the uncertainty in this estimate, however: in BM we showed, for example, that a model with flattening counts has much lower source confusion while remaining consistent with the observed counts at high flux densities. The actual confusion level remains to be determined from deeper counts at CMB frequencies (see Waldram et al. 2003, Waldram et al. 2004 for recent deep counts at 15 GHz).

Contamination in the multiband surveys is practically unaffected by point source confusion. For AMI we actually find a lower contamination rate, an apparent gain explained by the fact that the catalog now contains only the brighter SZ sources, due to the lowered sensitivity caused by point source confusion.

The photometry of the multiband surveys also shows little effect from the point sources. Fits to the recovered flux vs. true flux relation do not differ significantly from the no–source case, and the dispersion remains essentially the same. This is consistent with the idea that point source confusion is either modest compared to the noise (SPT) or controlled by multiband observations (Planck).

6. Discussion and Conclusion

We have described a simple, rapid method based on matched multi–frequency filters for extracting cluster catalogs from SZ surveys. We assessed its performance when applied to the three kinds of survey listed in Table 1. The rapidity of the method allows us to run many simulations of each survey to accurately quantify selection effects and observational uncertainties. We specifically examined catalog completeness, contamination rate and photometric precision.

Figure 2 shows the cluster selection criteria in terms of total SZ flux and source size. It clearly demonstrates that SZ surveys, in particular high resolution ground–bases surveys, will not be purely flux limited, something which must be correctly accounted for when interpreting catalog statistics (Melin et al. 2005).

Figure 3 and Table 2 summarize the expected yield for each survey. The counts roll off at the faint end well before the point–source flux limit (intercept of the curves in Figure 2 multiplied by the S/N limit) even at the high detection threshold of S/N=5; the surveys loose completeness precisely because they are not purely flux–limited. These yields depend on the underlying cluster model and are hence subject to non–negligible uncertainty. They are nonetheless indicative, and in this work we focus on the nature of observational selection effects for which the exact yields are of secondary importance.

At our fiducial S/N=5 detection threshold, overall catalog contamination remains below 5%, with some dependence on SZ flux for the single frequency survey (see Figure 5). The overall contamination rises to between 20%

and 30% at S/N>3. We note that the contamination rate is always larger than expected from pure instrumental noise, pointing to the influence of astrophysical confusion.

We pay particular attention to photometric precision, an issue often neglected in discussions of the scientific potential of SZ surveys. Scatter plots for the recovered flux for each survey type are given in Figures 7, 8 and 9. In the two multiband surveys, the recovered SZ flux is slightly biased, due to the flux cut, with a dispersion of $\sigma_{logY_\circ} = 0.17$ and $\sigma_{logY_\circ} = 0.13$ for SPT and Planck, respectively. This observational dispersion is significantly larger than the intrinsic dispersion in the $Y - M$ relation predicted by hydrodynamical simulations. This uncertainty must be properly accounted for in scientific interpretation of SZ catalogs; specifically, it will degrade survey calibration and cosmological constraints.

Even more importantly, we found that astrophysical confusion seriously compromises the photometry of the single frequency survey (Figure 9). The histogram in Figure 10 shows that the recovered flux has in fact a bimodal distribution. We traced the effect to an inability to determine source core radii in the presence of primary CMB anisotropy. If cluster core radius could be accurately measured, e.g., with X–ray follow–up, then we would obtain photometric precision comparable to the multiband surveys (see Figure 11). This confusion can also be removed by follow–up of detected sources at a second radio frequency (e.g., 90 GHz). Photometric uncertainty will therefore be key limiting factor in single frequency SZ surveys.

All these results apply to the ideal case where the filter exactly matches the (simulated) cluster profiles. We then examined the potential impact of cluster morphology and point sources on these conclusions.

Using N–body simulations, we found that cluster morphology has little effect on catalog completeness, but that it does increase the contamination rate and bias the photometry. The increased contamination is caused by deviations from a smooth radial SZ profile that appear as residual flux in the maps after source extraction. More importantly, the photometry is biased by the mismatch between the filter template and the actual cluster profile. This observational bias adds to the modeling uncertainty in the $Y - M$ relation, which will have to be empirically determined in order to use the catalog for cosmology studies.

As shown by Figure 13, point sources decrease survey completeness. The multiband surveys effectively reduce IR point source confusion and suffer only a small decrease. Radio source confusion, on the other hand, greatly decreased the completeness of the single frequency survey. This is consistent with the expectation that, for our adopted radio point source model and source subtraction threshold, point source confusion dominates instrumental noise. Modeling uncertainty here is, however, very large: radio source counts are not constrained at relevant fluxes ($\sim 100\,\mu$Jy), which requires us to extrapolate counts from mJy levels (see BM for a more detailed discussion).

Surveys based on the SZ effect will open a new window onto the high redshift universe. They inherit their strong scientific potential from the unique characteristics of the SZ signal. Full realization of this potential, however, requires understanding of observational selection effects and uncertainties. Overall, multiband surveys appear robust in this light, while single frequency surveys will most likely require additional observational effort, e.g., follow–up in other wavebands, to overcome large photometric errors caused by astrophysical confusion with primary CMB anisotropy.

References

Angulo R., Baugh C.M., Frenk C.S. et al. 2005, MNRAS, 362, L25

Arnaud M., astro–ph/0508159

Barbosa D., Bartlett J.G., Blanchard A., 1996, A&A, 314, 13

Bartelmann M. 2001, A&A, 370, 754

Birkinshaw M. 1999, Proc. 3K Cosmology, 476, American Institute of Physics, Woodbury, p. 298

Blanchard A. & Bartlett J.G. 1998, A&A 332, L49

Bartlett J. G. 2001, review in "Tracing cosmic evolution with galaxy clusters" (Sesto Pusteria 3-6 July 2001), ASP Conference Series in press, astro–ph/0111211

Bartlett J.G. & Melin, J.-B. 2005, submitted

Benson, A.J., Reichardt, C., Kamionkowski, M. 2002, MNRAS, 331, 71

Bertin E. & Arnouts S. 1996, A&A, 117, 393

Borys C., Chapman S., Halpern M. & Scott D. 2003, MNRAS, 344, 385

Carlstrom J. E., Holder G. P., Reese E. D. 2002, ARA&A, 40, 643

Colafrancesco, S., Mazzotta, P., Vittorio, N. 1997, ApJ, 488, 566

Cole S., Percival W.J., Peacock J.A. et al 2005, MNRAS, 362, 505

Delabrouille J., Melin J.-B., Bartlett J.G. 2002, in "AMiBA 2001: High-Z Clusters, Missing Baryons, and CMB Polarization", ASP Conference Proceedings, astro-ph/0109186

da Silva A.C., Kay S.T., Liddle A.R. & Thomas P.A. 2004, MNRAS, 348, 1401

Diego J.M., Vielva P., Martinez-Gonzalez E., Silk J., Sanz J.L. 2002, MNRAS, 336, 1351

Eisenstein D.J., Zehavi I., Hogg D.W. et al. 2005, ApJ, 633, 560

Eke V.R., Cole S., Frenk C.S. & Patrick Henry P.J. 1998, MNRAS, 298 1145

Freedman W.L., Madore B.F., Gibson B.K. *et al.* 2001, ApJ, 553, 47

Haehnelt, M. G. & Tegmark, M. 1996, MNRAS, 279, 545

Haiman Z., Mohr J.J. & Holder G.P. 2000, ApJ, 553, 545

Herranz D., Sanz J.L., Barreiro R.B., Martnez-Gonzalez E. 2002, ApJ, 580, 610

Herranz D., Sanz J.L., Hobson M.P., Barreiro R.B., Diego J.M., Martnez-Gonzalez E., Lasenby A. N. 2002, MNRAS, 336, 1057

Hobson M.P. & McLachlan C. 2003, MNRAS, 338, 765

Holder G. P., Mohr, J. J., Carlstrom J. E., Evrard A. E., Leitch E. M. 2000, ApJ, 544, 629

Hu W. 2003, Phys. Rev. D., 67, 081304

Huetsi G. 2005, A&A, 446, 43

Jenkins A., Frenk C.S., White S.D.M. *et al.* 2001, MNRAS, 321, 372

Jones M.E., Edge A.C., Grainge, K. *et al.* 2001, MNRAS, 357, 518

Juin J.B., Yvon D., Refregier A. & Yeche C. 2005, astro–ph/0512378

Kneissl, R., Jones, M.E., Saunders, R. *et al.* 2001, MNRAS, 328, 783

Knox L., Holder, G.P., Church S.E. 2004, ApJ, 612, 96

Lima M. & Hu W. 2004, Phys Rev D70, 043504

Lima M. & Hu W. 2005, Phys Rev D72, 043006

López-Caniego M., Herranz D., Sanz J.L., Barreiro R.B. 2005, astro-ph/0503149

Majumdar S. & Mohr J.J. 2004, ApJ 613, 41

Melin J.-B., Bartlett J.G., Delabrouille J. 2005, A & A 429, 417-426.

Mohr J. J., Mathiesen B., Evrard A. E. 1999, ApJ, 517, 627

Motl P.M., Hallman E.J., Burns J.O. & Norman M.L. 2005, ApJ, 623, L63

Nagai D., astro–ph/0512208

Oukbir J. & Blanchard A. 1992, A&A, 262, L21

Pierpaoli E., Anthoine S., Huffenberger K., Daubechies I. 2004, MNRAS, 359, 261

Pires S., Juin, J.B., Yvon D. et al., 0508641

Rosati, P., Borgani, S. & Norman C. 2002, ARA&A 40, 539

Ruhl J.E. et al. 2004, astro-ph/0411122

Schäfer B.M., Pfrommer C., Hell R.M, Bartelmann M. 2004, MNRAS, 370, 1713

Schulz, A. E., White, M. 2003, ApJ, 586, 723

Seljak U. & Zaldarriaga, M. 1996, ApJ, 469, 437, astro–ph/9603033, www.cmbfast.org

Spergel D. N. et al. 2003, ApJ Supplement Series, 148, 175

Spergel D. N. et al. 2006, astro–ph/0603449

Sunyaev, R. A. & Zel'dovich, Ya. B. 1970,Comments Astrophys. Space Phys., ComAp,2,66

Sunyaev, R. A. & Zel'dovich, Ya. B. 1972, Comments Astrophys. Space Phys., 4, 173

Vale C. & White M. 2005, New Astronomy, 11, 207

Voit G. M. 2004, astro–ph/0410173

Waldram E. M. et al. 2003, MNRAS 342, 915

Waldram E. M. & Pooley G. G. 2004, astro–ph/0407422

Wang S., Khoury J., Haiman, Z. & May M. 2004, PhysRev, D70, 12300

Weller J. & Battye R. A. 2003, NewAR, 47, 775

White M., 2003, ApJ, 597, 650